

```

seq_documentation_block:
LOCUS BE839427 564 bp mRNA EST 22-SEP-2000
DEFINITION RC3-FN0143-190700-022-d12 FN0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE839427
VERSION BE839427.1 GI:10271805
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baig,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC3-FN0143-190
700-022-d12&t3=2000-07-19&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 564.
Location/Qualifiers
1..564
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0143"
/dev_stage="Adult"
/note="organ: prostate_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
108 a 156 c 121 g 179 t

```

## FEATURES

source

```

seq_documentation_block:
LOCUS BE523624 538 bp mRNA EST 11-DEC-2000
DEFINITION UI-R-Cl-1c-d-10-0-UI-r1 UI-R-Cl Rattus norvegicus cDNA clone
UI-R-Cl-1c-d-10-0-UI 5', mRNA sequence.
ACCESSION BE523624
VERSION BE523624.1 GI:11631591
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 538)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
This clone is also available through Research Genetics (www.resgen.com)
LNL (info@image.lnl.gov). IMAGE ID- 1792594
Seq primer: M13 Forward.
Location/Qualifiers
1..538
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Cl-1c-d-10-0-UI"

```

## BASE ORIGIN

```

alignment_scores:
Quality: 883.00 Length: 175
Ratio: 5.075 Gaps: 0
Percent Similarity: 99.429 Percent Identity: 98.857

```

## alignment\_block:

US-09-513-365A-1 x BE839427/rev ..

Align seg 1/1 to reverse of: BE839427 from: 1 to: 564

```

1 MetLeuGlyClnGlnGlnGlnLeuTyrSerSerAlaAlaLeuTh 17
|||||
534 ATGTAGGAGCAGCAGCAGCAACTACTCTCGCGCTCCGTCGAC 485
17 fGlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluC 34
|||||
484 CGGGAGCGGAGCGGCTGCTACCTGCTACGTGCAGGACTACCTTAGT 435
34 ysValGluSerLeuProHisAspMetGlnArgAsnValSerValLeuArg 50

```

## FEATURES

source

```

1..538
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Cl-1c-d-10-0-UI"

```

```

434 GCGTGGAGTCGCTCCCGACACATGCAGAGAACGTCCTCTGCTGCGA 385
51 GluLeuAspAsnLysTyrGlnGlnThrLeuLysGlnLysAspValTy 67
384 GAGCTGGACAAACAATATCAAGAAACGTTAAAGGAATTTGATGATGCTA 335
67 xGluLysTyrLysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnG 84
334 CGAAAAATATTAAGAAAGAGATGATTTAAACCAAGAAACGTCCTACAGC 285
84 InLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlnGlnGlnGlnGln 100
284 AGCTTCCAGAGACACTAATTAATAGTCAAGATTTGGAGATGAAGAA 235
101 IleGlnIleValThrGlnMetLeuGlnValGlnAsnArgAlaArgG 117
234 ATACAGATTGTTACACAATGCTCGAATTTGTTGGAATCGGCAAGACA 185
117 nMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArg 134
184 AATGGAGTTACACTCACAGTGTTCCTCAAGATCCTGCTGAAAGTGAACGAG 135
134 laSerAspLysAlaLysMetAspSerSerGlnProGlnArgSerSerArg 150
134 CCTCAGATAAGCAAGATGATTCACCAACCAAGAAAGATCTTCAAGA 85
151 ArgProArgGlnArgThrSerGluSerArgAspLeuCysHisMetal 167
84 AGACCCCGCAGCGACGCGACAGTGTAAAGCGGTGATTATGTACATGCG 35
167 aAsnGlyIleGluAspCysAspAsp 175
34 AAATGGGATTGAAGACTGTGATGAG 10

```

seq\_name: gb\_est89:BF523624

seq\_documentation\_block:

```

LOCUS BF523624 538 bp mRNA EST 11-DEC-2000
DEFINITION UI-R-Cl-1c-d-10-0-UI-r1 UI-R-Cl Rattus norvegicus cDNA clone
UI-R-Cl-1c-d-10-0-UI 5', mRNA sequence.
ACCESSION BF523624
VERSION BF523624.1 GI:11631591
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 538)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
This clone is also available through Research Genetics (www.resgen.com)
LNL (info@image.lnl.gov). IMAGE ID- 1792594
Seq primer: M13 Forward.
Location/Qualifiers
1..538
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Cl-1c-d-10-0-UI"

```

/clone\_lib="UI-R-C1"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="vector: p773D-pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C1  
 library is a subtracted library derived from the UI-R-C0  
 library, which is a subtracted library derived from the  
 UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library  
 consisted of a mixture of individually tagged normalized  
 libraries constructed from rat placenta, adult lung,  
 brain, liver, kidney, heart, spleen, ovary, and muscle.  
 The UI-R-E1 library consisted of a mixture of  
 individually tagged normalized libraries constructed from  
 8, 12 and 18-day embryo. The tag is a string of 3-5  
 nucleotides present between the Not I site and the  
 oligo-dT track which allows identification of the library  
 of origin of a clone within the mixture. The subtracted  
 library (UI-R-C1) was constructed as follows: PCR  
 amplified cDNA inserts from UI-R-C0 clones from which 3'  
 ESTs had been derived was used as a driver in a  
 hybridization with the UI-R-C0 library in the form of  
 single-stranded circles. The remaining single-stranded  
 circles (subtracted library) was purified by  
 hydroxyapatite column chromatography, converted to  
 double-stranded circles and electroporated into DH10B  
 bacteria (Life Technologies) to generate the UI-R-C1  
 library. This procedure has been previously described  
 (Bonaldo, Lennon and Soares, Genome Research 5: 791-806,  
 1996)."

BASE COUNT 95 a 135 c 117 g 191 t  
 ORIGIN

alignment\_scores:  
 Quality: 874.00 Length: 175  
 Ratio: 5.052 Gaps: 0  
 Percent Similarity: 98.857 Percent Identity: 96.571

alignment\_block:  
 US-09-513-365A-1 x BF523624/rev ..

Align seg 1/1 to reverse of: BF523624 from: 1 to: 538

```

17 ThrGlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuG1 33
|||||
536 ACCGAGAGCGAGCGGCTCCTACCTGCTACGTGCAGGACTACTGGA 487
|||||
33 ucysValGluSerLeuProHisAspMetGlnArgAsnValSerValLeuA 50
|||||
336 GTGTGTGGAGTCGTGCCCGCCACGACATGCAGAGGAACGTGCTGCTGC 437
|||||
50 rgGluLeuAspAsnLysTyrGlnGluThrLeuLysGluLeuAspVal 66
|||||
436 GGGAGCTGGACACACATACCCAGAGAACTTTAAGGAATGATGATGTC 387
|||||
67 TyrGluLysTyrLysLysGluAspLeuAsnGlnLysLysArgLeuG1 83
|||||
386 TATCAAAAATATAAGAAAGAGATGATTCAACACGAAAAACGCTTACA 337
|||||
83 nGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGluL 100
|||||
336 GCACATCTCCAGAGAGCATTAATCAATAGCAAGATTTGGAGATGAAA 287
|||||
100 ysileGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArg 116
|||||
286 AGATTGAGATTGTCCACAGATGCTGGAAATTTGGTGGAAACCGACCCA 237
|||||
117 GlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluAr 133
|||||
236 CAATGGAGCTGCATTACAGTGTCTCCAGATCTCTGCTGAAGTGACGC 187
|||||
133 gAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSera 150
|||||

```

```

186 GGCTCAGACAAATCCAAAGATGGATTCCAGTCAACCGAAGATCTTCCA 137
150 rgArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMet 166
|||||
136 GRAGACTCGAAGACAGCGGACGAGTGGAGCCGCGTATTATGTCACATG 87
|||||
167 AlaAsnGlyIleGluAspCysAspAspGlnProProlLysGluLysLysSe 183
:::|||||
86 ACAAGCGGATGTCGATGCTGATGATGATGATGATGATGATGATGATGAT 37
183 rLysSerAlaLysLysLysLysArg 191
|||||
36 CAATCCGCGCAAAAAAAGAAAAA 12
seq_name: gb_est79:BE839460

seq_documentation_block:
LOCUS BE839460 566 bp mRNA EST 22-SEP-2000
DEFINITION RC3-FN0143-260700-012-cl2 FN0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE839460
VERSION BE839460.1 GI:10271838
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 566)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=cl2-RC3-FN0143-260
700-012-cl2&t3=2000-07-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 565.
FEATURES
Location/Qualifiers
source
1..566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0143"
/dev_stage="Adult"
/note="Organ: prostate.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 177 a 124 c 160 g 105 t
ORIGIN

alignment_scores:
Quality: 852.00 Length: 169
Ratio: 5.041 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.817

```

5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15

```

|||||
142 CGGGGAGCGGAGCGGCTGCTACCTGCTACGTCAGGACTACCTGAGT 191
34 yValGluSerLeuProHisAspMetGlnArgasnValSerValLeuArg 50
|||||
192 GCGTGAGTGCCTGCCCCAGCATGCAGAGAACGTCGTCTGTCTGCGA 241
51 GluLeuAspAsnLysTyrGlnGluThrLeuLysGluLeuLeuAspVal 67
|||||
242 GAGCTGGACACAAATATCAAGAACGTTAAAGAAATGATGATGCTA 291
67 rGluLysTyrLysLysGluAspAspLeuAsnGlnLysLysArgLeuGln 84
|||||
292 CGAAAAATATAAGAAAGAGATGATTAACACAGAGAACGCTACAGC 341
84 lneLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGluLys 100
|||||
342 AGCTTCTCCAGAGAGACATAATTAATAGTCAAGAAATGGGAGATGAAA 391
101 lIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArg 117
|||||
2 ATACAGATTGTTACACAAATGCTCGAATGGTGGAAAAATCGGCAAGACA 441
117 nMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArg 134
|||||
442 AATGAGATTACACTACAGTGTTCACAGATCCTGCTGAAAGTGAACGAG 491
134 lAserAspLysAlaLysMetAspSerSerGlnProGluArgSerArg 150
|||||
492 CTCAGATAAGCAAGATGGATTTCAGCCCAACAGAAAGATCTTCAAGA 541
151 ArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAl 167
|||||
542 AGACCCCGCAGCAGCGGACAGTGAAGCCGTGATTATGTCACATGGC 591
167 asnGlylIleGluAspCysAspAspGlnProProLysGluLysSerL 184
|||||
592 AAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAAAGAAATCCA 641
184 ySerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
|||||
642 AGTCAGCAAGAAAAAGAACGCTCCAGGCCAAGCAGGAAGGGAAGCT 691
201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCy 217
|||||
692 TCACCTGTTGAGTTGCAATAGATCTCTAATGAACCTACATAGCTATG 741
217 shsGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysP 234
|||||
2 CAACCAAGTGTCTTATGGGAGATGATAGATGTGACAAATGAACAGTGC 791
234 toIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLys 250
|||||
792 CAATTGAATGGTTTCACTTTTCAATGTTTCACTTACCTATAAACCAAG 841
251 GlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThrMetAs 267
|||||
842 GGGAAATGATATGCCCCAAGTCCAGGGGAGATATGACAAAACAATGGA 891
267 pLysSerThrGluLysThrLysLysAspArgArgSerArg 280
|||||
892 CAAAAAGTACTGAAAAACAAAAAAGGATAGAAAGATCGAGG 931

```

seq\_name: gb\_pr4:AF053537

seq\_documentation\_block:

LOCUS AF053537 1080 bp mRNA PRI 08-SEP-2000  
 DEFINITION Homo sapiens p33 (ING2) mRNA, complete cds.  
 ACCESSION AF053537  
 VERSION AF053537.1 GI:9992837

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

6

1

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1080)  
 Nagashima, M., Hagiwara, K., Minter, A.R. and Harris, C.C.  
 Direct Submission  
 Submitted (12-MAR-1998) Laboratory of Human Carcinogenesis,  
 National Cancer Institute, 37 Convent Drive Bldg. 37 Rm. 2C01,  
 Bethesda, MD 20892, USA

FEATURES  
 Location/Qualifiers  
 1..1080

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4q35"  
 1..1080  
 /gene="ING2"  
 68..910  
 /gene="ING2"  
 /codon\_start=1  
 /product="p33"  
 /protein\_id="AAG11395.1"  
 /db\_xref="GI:9992838"  
 /translation="MLGQOQOOLYSAAALLTGPERSLLTCYVODYLECVESLPHDMQR  
 NVSLRELNDKYQETLKEIDDDYKYEKKEDDLNKKRLQQLLORALINSDELGDKIO  
 IVTQMLELVENRARMELHSCQFQDPAESERASDKAKMDSSQPSRRSRPRORTSES  
 RDLCHWANGIEDCDOPPKKKSKSAKKKRSKAKOERASPVFEAIDNPETCYLGN  
 QVSYGEMIGCDNEQCPIEWHFSCVSLTYKPKGKWCYCPKCRGDNKMTDKSTERTKTD  
 RRSR"  
 106  
 /gene="ING2"  
 /note="polymorphism"  
 /replace="c"  
 375 a 196 c 271 g 238 t

variation

BASE COUNT 375 a 196 c 271 g 238 t  
 ORIGIN

alignment\_scores:  
 Quality: 1481.00 Length: 280  
 Ratio: 5.289 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-513-365A-1 x AF053537 ..  
 Align seg 1/1 to: AF053537 from: 1 to: 1080

1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuTh 17  
 |||||  
 68 ATGTTAGGGCAGCAGCAGCAGCACTACTCTCGCTCGCTCGCTCTGAC 117  
 |||||  
 17 rGlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGlu 34  
 |||||  
 118 CGGGGAGCGGAGCGGCTGCTACCTGCTACGTCAGGACTACCTTGAGT 167  
 |||||  
 34 yValGluSerLeuProHisAspMetGlnArgasnValSerValLeuArg 50  
 |||||  
 168 GCGTGAGTGCCTGCCCCAGCATGCAGAGAACGTCGTCTGTCTGCGA 217  
 |||||  
 51 GluLeuAspAsnLysTyrGlnGluThrLeuLysGluLeuAspVal 67  
 |||||  
 218 GAGCTGGACACAAATATCAAGAACGTTAAAGAAATGATGATGCTA 267  
 |||||  
 67 rGluLysTyrLysLysGluAspAspLeuAsnGlnLysLysArgLeuGln 84  
 |||||  
 268 CGAAAAATATAAGAAAGAGATGATTAACACAGAGAACGCTACAGC 317  
 |||||  
 84 lneLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGluLys 100  
 |||||  
 318 AGCTTCTCCAGAGAGACATAATTAATAGTCAAGAAATGGGAGATGAAA 367  
 |||||  
 101 lIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArg 117  
 |||||  
 368 ATACAGATTGTTACACAAATGCTCGAATGGTGGAAAAATCGGCAAGACA 417